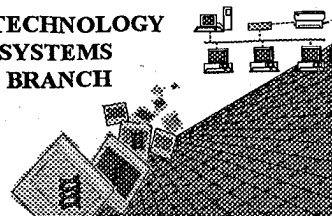


05-90
0328

BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/001,486
Source: Q1PE
Date Processed by STIC: 3/25/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10/001,486

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☒ Wrapped Nucleics
Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 ☐ Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 ☒ Misaligned Amino
Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 ☐ Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 ☐ Variable Length
Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 ☐ PatentIn 2.0
"bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s). Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 ☐ Skipped Sequences
(OLD RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 ☐ Skipped Sequences
(NEW RULES)
Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 ☐ Use of n's or Xaa's
(NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ☐ Invalid <213>
Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ☐ Use of <220>
Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 ☐ PatentIn 2.0
"bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 ☐ Misuse of n
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002
 TIME: 14:48:40

Input Set : A:\W065457.txt
 Output Set: N:\CRF3\03252002\J001486.raw

Does Not Comply
 Corrected Diskette Needed

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
 6 (i) APPLICANT: The Government of the United States of America
 7 *delete's* as represented by the Secretary
 8 Department of Health and Human Services
 9 Washington, D.C.
 10 Htun Ph.D., Han
 11 Hager Ph.D., Gordon L.
 14 (ii) TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
 15 DNA BINDING MOLECULES IN LIVING CELLS
 17 (iii) NUMBER OF SEQUENCES: 9
 19 (iv) CORRESPONDENCE ADDRESS:
 20 (A) ADDRESSEE: Needle & Rosenberg
 21 (B) STREET: 127 Peachtree Street, Suite 1200
 22 (C) CITY: Atlanta
 23 (D) STATE: Georgia
 24 (E) COUNTRY: USA
 25 (F) ZIP: 30303
 27 (v) COMPUTER READABLE FORM:
 28 (A) MEDIUM TYPE: Floppy disk
 29 (B) COMPUTER: IBM PC compatible
 30 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 31 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 33 (vi) CURRENT APPLICATION DATA:
 34 (A) APPLICATION NUMBER: US/10/001,486
 35 (B) FILING DATE: 11-Mar-2002
 41 (C) CLASSIFICATION:
 38 *ote* (vii) PRIOR APPLICATION DATA:
 39 (A) APPLICATION NUMBER: 60/008,373
 40 (B) FILING DATE: 08 Dec 1995
 43 (viii) ATTORNEY/AGENT INFORMATION:
 44 (A) NAME: Selby, Elizabeth
 45 (B) REGISTRATION NUMBER: 38298
 46 (C) REFERENCE/DOCKET NUMBER: 14014.0183
 48 (ix) TELECOMMUNICATION INFORMATION:
 49 (A) TELEPHONE: 404-688-0770
 50 (B) TELEFAX: 404-688-9880

ERRORED SEQUENCES

53 (2) INFORMATION FOR SEQ ID NO: 1:
 55 (i) SEQUENCE CHARACTERISTICS:

ppr 2-10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

56      (A) LENGTH: 7257 base pairs
57      (B) TYPE: nucleic acid
58      (C) STRANDEDNESS: double
C--> 59      (D) TOPOLOGY: linear
W--> 61      (ii) MOLECULE TYPE: nucleic acid
64      (ix) FEATURE:
65          (A) NAME/KEY: CDS
66          (B) LOCATION: 1072..4284
68      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
E--> 70 TCAATATTGG CCATTAGCCA TATTATTCAT TGGTTATATA GCATAAATCA
71 ATATTGGCTA 60
E--> 73 TTGGCCATTG CATACGTTGT ATCTATATCA TAATATGTAC ATTTATATTG
74 GCTCATGTCC 120
E--> 76 AATATGACCG CCATGTTGGC ATTGATTATT GACTAGTTAT TAATAGTAAT
77 CAATTACGGG 180
E--> 79 GTCATTAGTT CATAGCCCAT ATATGGAGTT CCGCGTTACA TAACTTACGG
80 TAAATGGCCC 240
E--> 82 GCCTGGCTGA CCGCCCAACG ACCCCCGCCC ATTGACGTCA ATAATGACGT
83 ATGTTCCCAT 300
E--> 85 AGTAACGCCA ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTTAC
86 GGTAAACTGC 360
E--> 88 CCACTTGGCA GTACATCAAG TGTATCATAT GCCAAGTCCG CCCCCTATTG
89 ACGTCAATGA 420
E--> 91 CGGTAAATGG CCCGCCTGGC ATTATGCCCA GTACATGACC TTACGGGACT
92 TTCCTACTTG 480
E--> 94 GCAGTACATC TACGTATTAG TCATCGCTAT TACCATGGTG ATGCGGTTTT
95 GGCAGTACAC 540
E--> 97 CAATGGGCGT GGATAGCGGT TTGACTCAGC GGGATTTCCTA AGTCTCCACC
98 CCATTGACGT 600
E--> 100 CAATGGGAGT TTGTTTGGC ACCAAAATCA ACGGGACTTT CCAAAATGTC
101 GTAATAACCC 660
E--> 103 CGCCCCGTTG ACGCAAATGG GCGGTAGGCG TGTACGGTGG GAGGTCTATA
104 TAAGCAGAGC 720
E--> 106 TCGTTTAGTG AACCCTCAGA TCACTAGAAG CTTTATTGCG GTAGTTTATC
107 ACAGTTAAAT 780
E--> 109 TGCTAACGCA GTCAGTGCTT CTGACACAAC AGTCTCGAAC TTAAGCTGCA
110 GAAGTTGGTC 840
E--> 112 GTGAGGCACT GGGCAGGTAA GSTATCAAGG TACAAGACAG GTTTAAGGAG
113 ACCAATAGAA 900
E--> 115 ACTGGGCTTG TCGAGACAGA GAAGACTCTT GCGTTTCTGA TAGGCACCTA
116 TTGGTCTTAC 960
E--> 118 TGACATCCAC TTTGCCTTTC TCTCCACAGG TGTCCACTCC CAGTTCAATT
119 ACAGCTCTTA 1020
E--> 121 AGGCTAGAGT ACTTAATACG ACTCACTATA GGCTAGCGAA GGAGATCCGC C ATG
122 GCC 1077
123
124
E--> 126 CAC CAT CAC CAC CAT CAC GGA TAT CCA TAC GAC GTG CCA GAT TAC GCT
W--> 127 1125

```

*global**format error**(see item 1
on Error**Summary Sheet)**same
error*

Met Ala

1

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

128 His His His His His His Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
W--> 129 5 10 15
E--> 131 CAG TCG AGT GCC ATG AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT GTC
W--> 132 1173
133 Gln Ser Ser Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val
W--> 134 20 25 30
E--> 137 CCA ATT CTT GTT GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT TCT
W--> 138 1221
139 Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser
W--> 140 35 40 45 50
E--> 142 GTC AGT GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC CTT
W--> 143 1269
144 Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu
W--> 145 55 60 65
E--> 147 AAA TTT ATT TGC ACT ACT GGA AAA CTA CCT GTT CCT TGG CCA ACA CTT
W--> 148 1317
149 Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu
W--> 150 70 75 80
E--> 152 GTC ACT ACT TTC ACT TAT GGT GTT CAA TGC TTT TCA AGA TAC CCA GAT
W--> 153 1365
154 Val Thr Thr Phe Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp
W--> 155 85 90 95
E--> 157 CAT ATG AAA CAG CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT
W--> 158 1413
159 His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr
W--> 160 100 105 110
E--> 162 GTA CAG GAA AGA ACT ATA TTT TTC AAA GAT GAC GGG AAC TAC AAG ACA
W--> 163 1461
164 Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr
W--> 165 115 120 125 130
E--> 167 CGT GCT GAA GTC AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG
W--> 168 1509
169 Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu
W--> 170 135 140 145
E--> 172 TTA AAA GGT ATT GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAC AAA
W--> 173 1557
174 Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys
W--> 175 150 155 160
E--> 177 TTG GAA TAC AAC TAT AAC TCA CAC AAT GTA TAC ATC ATG GCA GAC AAA
W--> 178 1605
179 Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys
W--> 180 165 170 175
E--> 182 CAA AAG AAT GGA ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT GAA
W--> 183 1653
184 Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu
W--> 185 180 185 190
E--> 187 GAT GGA AGC GTT CAA CTA GCA GAC CAT TAT CAA CAA AAT ACT CCA ATT
W--> 188 1701
189 Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile

*same
error*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

W--> 190 195                200                205                210
E--> 192 GGC GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA CAA
W--> 193 1749
      194 Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln
W--> 195                215                220                225
E--> 197 TCT GCC CTT TCG AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC CTT
W--> 198 1797
      199 Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu
W--> 200                230                235                240
E--> 202 CTT GAG TTT GTA ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA
W--> 203 1845
      204 Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu
W--> 205                245                250                255
E--> 209 TAC AAA GGC GCC GGC GCT GGT GCT GGT GCT GGC GCC ATC AGC GCG CTG
W--> 210 1893
      211 Tyr Lys Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Ile Ser Ala Leu
W--> 212                260                265                270
E--> 214 ATC CTG GAC TCC AAA GAA TCC TTA GCT CCC CCT GGT AGA GAC GAA GTC
W--> 215 1941
      216 Ile Leu Asp Ser Lys Glu Ser Leu Ala Pro Pro Gly Arg Asp Glu Val
W--> 217 275                280                285                290
E--> 219 CCT GGC AGT TTG CTT GGC CAG GGG AGG GGG AGC GTA ATG GAC TTT TAT
W--> 220 1989
      221 Pro Gly Ser Leu Leu Gly Gln Gly Arg Gly Ser Val Met Asp Phe Tyr
W--> 222                295                300                305
E--> 224 AAA AGC CTG AGG GGA GGA GCT ACA GTC AAG GTT TCT GCA TCT TCG CCC
W--> 225 2037
      226 Lys Ser Leu Arg Gly Gly Ala Thr Val Lys Val Ser Ala Ser Ser Pro
W--> 227                310                315                320
E--> 229 TCA GTG GCT GCT GCT TCT CAG GCA GAT TCC AAG CAG CAG AGG ATT CTC
W--> 230 2085
      231 Ser Val Ala Ala Ala Ser Gln Ala Asp Ser Lys Gln Gln Arg Ile Leu
W--> 232                325                330                335
E--> 234 CTT GAT TTC TCG AAA GGC TCC ACA AGC AAT GTG CAG CAG CGA CAG CAG
W--> 235 2133
      236 Leu Asp Phe Ser Lys Gly Ser Thr Ser Asn Val Gln Gln Arg Gln Gln
W--> 237                340                345                350
E--> 239 CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG
W--> 240 2181
      241 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
W--> 242 355                360                365                370
E--> 244 CAG CCA GGC TTA TCC AAA GCC GTT TCA CTG TCC ATG GGG CTG TAT ATG
W--> 245 2229
      246 Gln Pro Gly Leu Ser Lys Ala Val Ser Leu Ser Met Gly Leu Tyr Met
W--> 247                375                380                385
E--> 249 GGA GAG ACA GAA ACA AAA GTG ATG GGG AAT GAC TTG GGC TAC CCA CAG
W--> 250 2277
      251 Gly Glu Thr Glu Thr Lys Val Met Gly Asn Asp Leu Gly Tyr Pro Gln
W--> 252                390                395                400

```

same

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

E--> 254 CAG GGC CAA CTT GGC CTT TCC TCT GGG GAA ACA GAC TTT CGG CTT CTG
W--> 255 2325
256 Gln Gly Gln Leu Gly Leu Ser Ser Gly Glu Thr Asp Phe Arg Leu Leu
W--> 257 405 410 415
E--> 259 GAA GAA AGC ATT GCA AAC CTC AAT AGG TCG ACC AGC GTT CCA GAG AAC
W--> 260 2373
261 Glu Glu Ser Ile Ala Asn Leu Asn Arg Ser Thr Ser Val Pro Glu Asn
W--> 262 420 425 430
E--> 264 CCC AAG AGT TCA ACG TCT GCA ACT GGG TGT GCT ACC CCG ACA GAG AAG
W--> 265 2421
266 Pro Lys Ser Ser Thr Ser Ala Thr Gly Cys Ala Thr Pro Thr Glu Lys
W--> 267 435 440 445 450
E--> 269 GAG TTT CCC AAA ACT CAC TCG GAT GCA TCT TCA GAA CAG CAA AAT CGA
W--> 270 2469
271 Glu Phe Pro Lys Thr His Ser Asp Ala Ser Ser Glu Gln Gln Asn Arg
W--> 272 455 460 465
E--> 274 AAA AGC CAG ACC GGC ACC AAC GGA GGC AGT GTG AAA TTG TAT CCC ACA
W--> 275 2517
276 Lys Ser Gln Thr Gly Thr Asn Gly Gly Ser Val Lys Leu Tyr Pro Thr
W--> 277 470 475 480
E--> 281 GAC CAA AGC ACC TTT GAC CTC TTG AAG GAT TTG GAG TTT TCC GCT GGG
W--> 282 2565
283 Asp Gln Ser Thr Phe Asp Leu Leu Lys Asp Leu Glu Phe Ser Ala Gly
W--> 284 485 490 495
E--> 286 TCC CCA AGT AAA GAC ACA AAC GAG AGT CCC TGG AGA TCA GAT CTG TTG
W--> 287 2613
288 Ser Pro Ser Lys Asp Thr Asn Glu Ser Pro Trp Arg Ser Asp Leu Leu
W--> 289 500 505 510
E--> 291 ATA GAT GAA AAC TTG CTT TCT CCT TTG GCG GGA GAA GAT GAT CCA TTC
W--> 292 2661
293 Ile Asp Glu Asn Leu Leu Ser Pro Leu Ala Gly Glu Asp Asp Pro Phe
W--> 294 515 520 525 530
E--> 296 CTT CTC GAA GGG AAC ACG AAT GAG GAT TGT AAG CCT CTT ATT TTA CCG
W--> 297 2709
298 Leu Leu Glu Gly Asn Thr Asn Glu Asp Cys Lys Pro Leu Ile Leu Pro
W--> 299 535 540 545
E--> 301 GAC ACT AAA CCT AAA ATT AAG GAT ACT GGA GAT ACA ATC TTA TCA AGT
W--> 302 2757
303 Asp Thr Lys Pro Lys Ile Lys Asp Thr Gly Asp Thr Ile Leu Ser Ser
W--> 304 550 555 560
E--> 306 CCC AGC AGT GTG GCA CTA CCC CAA GTG AAA ACA GAA AAA GAT GAT TTC
W--> 307 2805
308 Pro Ser Ser Val Ala Leu Pro Gln Val Lys Thr Glu Lys Asp Asp Phe
W--> 309 565 570 575
E--> 311 ATT GAA CTT TGC ACC CCC GGG GTA ATT AAG CAA GAG AAA CTG GGC CCA
W--> 312 2853
313 Ile Glu Leu Cys Thr Pro Gly Val Ile Lys Gln Glu Lys Leu Gly Pro
W--> 314 580 585 590
E--> 316 GTT TAT TGT CAG GCA AGC TTT TCT GGG ACA AAT ATA ATT GGT AAT AAA

same

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

W--> 317 2901
 318 Val Tyr Cys Gln Ala Ser Phe Ser Gly Thr Asn Ile Ile Gly Asn Lys
 W--> 319 595 600 605 610
 E--> 321 ATG TCT GCC ATT TCT GTT CAT GGT GTG AGT ACC TCT GGA GGA CAG ATG
 W--> 322 2949
 323 Met Ser Ala Ile Ser Val His Gly Val Ser Thr Ser Gly Gly Gln Met
 W--> 324 615 620 625
 E--> 326 TAC CAC TAT GAC ATG AAT ACA GCA TCC CTT TCT CAG CAG CAG GAT CAG
 W--> 327 2997
 328 Tyr His Tyr Asp Met Asn Thr Ala Ser Leu Ser Gln Gln Gln Asp Gln
 W--> 329 630 635 640
 E--> 331 AAG CCT GTT TTT AAT GTC ATT CCA CCA ATT CCT GTT GGT TCT GAA AAC
 W--> 332 3045
 333 Lys Pro Val Phe Asn Val Ile Pro Pro Ile Pro Val Gly Ser Glu Asn
 W--> 334 645 650 655
 E--> 336 TGG AAT AGG TGC CAA GGC TCC GGA GAG GAC AGC CTG ACT TCC TTG GGG
 W--> 337 3093
 338 Trp Asn Arg Cys Gln Gly Ser Gly Glu Asp Ser Leu Thr Ser Leu Gly
 W--> 339 660 665 670
 E--> 341 GCT CTG AAC TTC CCA GGC CGG TCA GTG TTT TCT AAT GGG TAC TCA AGC
 W--> 342 3141
 343 Ala Leu Asn Phe Pro Gly Arg Ser Val Phe Ser Asn Gly Tyr Ser Ser
 W--> 344 675 680 685 690
 E--> 346 CCT GGA ATG AGA CCA GAT GTA AGC TCT CCT CCA TCC AGC TCG TCA GCA
 W--> 347 3189
 348 Pro Gly Met Arg Pro Asp Val Ser Ser Pro Pro Ser Ser Ser Ala
 W--> 349 695 700 705
 E--> 353 GCC ACG GGA CCA CCT CCC AAG CTC TGC CTG GTG TGC TCC GAT GAA GCT
 W--> 354 3237
 355 Ala Thr Gly Pro Pro Pro Lys Leu Cys Leu Val Cys Ser Asp Glu Ala
 W--> 356 710 715 720
 E--> 358 TCA GGA TGT CAT TAC GGG GTG CTG ACA TGT GGA AGC TGC AAA GTA TTC
 W--> 359 3285
 360 Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys Val Phe
 W--> 361 725 730 735
 E--> 363 TTT AAA AGA GCA GTG GAA GGA CAG CAC AAT TAC CTT TGT GCT GGA AGA
 W--> 364 3333
 365 Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala Gly Arg
 W--> 366 740 745 750
 E--> 368 AAC GAT TGC ATC ATT GAT AAA ATT CGA AGG AAA AAC TGC CCA GCA TGC
 W--> 369 3381
 370 Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro Ala Cys
 W--> 371 755 760 765 770
 E--> 373 CGC TAT CGG AAA TGT CTT CAG GCT GGA ATG AAC CTT GAA GCT CGA AAA
 W--> 374 3429
 375 Arg Tyr Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys
 W--> 376 775 780 785
 E--> 378 ACA AAG AAA AAA ATC AAA GGG ATT CAG CAA GCC ACT GCA GGA GTC TCA
 W--> 379 3477

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

      380 Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Ala Thr Ala Gly Val Ser
W--> 381              790              795              800
E--> 383 CAA GAC ACT TCG GAA AAT CCT AAC AAA ACA ATA GTT CCT GCA GCA TTA
W--> 384 3525
      385 Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Ile Val Pro Ala Ala Leu
W--> 386              805              810              815
E--> 388 CCA CAG CTC ACC CCT ACC TTG GTG TCA CTG CTG GAG GTG ATT GAA CCC
W--> 389 3573
      390 Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Leu Glu Val Ile Glu Pro
W--> 391              820              825              830
E--> 393 GAG GTG TTG TAT GCA GGA TAT GAT AGC TCT GTT CCA GAT TCA GCA TGG
W--> 394 3621
      395 Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Val Pro Asp Ser Ala Trp
W--> 396 835              840              845              850
E--> 398 AGA ATT ATG ACC ACA CTC AAC ATG TTA GGT GGG CGT CAA GTG ATT GCA
W--> 399 3669
      400 Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gly Arg Gln Val Ile Ala
W--> 401              855              860              865
E--> 403 GCA GTG AAA TGG GCA AAG GCG ATA CTA GGC TTG AGA AAC TTA CAC CTC
W--> 404 3717
      405 Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Leu Arg Asn Leu His Leu
W--> 406              870              875              880
E--> 408 GAT GAC CAA ATG ACC CTG CTA CAG TAC TCA TGG ATG TTT CTC ATG GCA
W--> 409 3765
      410 Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Trp Met Phe Leu Met Ala
W--> 411              885              890              895
E--> 413 TTT GCC TTG GGT TGG AGA TCA TAC AGA CAA TCA AGC GGA AAC CTG CTC
W--> 414 3813
      415 Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Ser Ser Gly Asn Leu Leu
W--> 416          900              905              910
E--> 418 TGC TTT GCT CCT GAT CTG ATT ATT AAT GAG CAG AGA ATG TCT CTA CCC
W--> 419 3861
      420 Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gln Arg Met Ser Leu Pro
W--> 421 915              920              925              930
E--> 425 GGC ATG TAT GAC CAA TGT AAA CAC ATG CTG TTT GTC TCC TCT GAA TTA
W--> 426 3909
      427 Gly Met Tyr Asp Gln Cys Lys His Met Leu Phe Val Ser Ser Glu Leu
W--> 428              935              940              945
E--> 430 CAA AGA TTG CAG GTA TCC TAT GAA GAG TAT CTC TGT ATG AAA ACC TTA
W--> 431 3957
      432 Gln Arg Leu Gln Val Ser Tyr Glu Glu Tyr Leu Cys Met Lys Thr Leu
W--> 433              950              955              960
E--> 435 CTG CTT CTC TCC TCA GTT CCT AAG GAA GGT CTG AAG AGC CAA GAG TTA
W--> 436 4005
      437 Leu Leu Leu Ser Ser Val Pro Lys Glu Gly Leu Lys Ser Gln Glu Leu
W--> 438              965              970              975
E--> 440 TTT GAT GAG ATT CGA ATG ACT TAT ATC AAA GAG CTA GGA AAA GCC ATC
W--> 441 4053
      442 Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Glu Leu Gly Lys Ala Ile

```

same

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

E--> 513 GCGGTATTTT CTCCTTACGC ATCTGTGCGG TATTTACACAC CGCATATGGT
514 GCACTCTCAG 5194
E--> 516 TACAATCTGC TCTGATGCCG CATAGTTAAG CCAGCCCCGA CACCCGCCAA
517 CACCCGCTGA 5254
E--> 519 CGCGCCCTGA CGGGCTTGTC TGCTCCCGGC ATCCGCTTAC AGACAAGCTG
520 TGACCGTCTC 5314
E--> 522 CGGGAGCTGC ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCGCGA
523 GACGAAAGGG 5374
E--> 525 CCTCGTGATA CGCCTATTTT TATAGGTTAA TGTATGATA ATAATGGTTT
526 CTTAGACGTC 5434
E--> 528 AGGTGGCACT TTTCCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT
529 TCTAAATACA 5494
E--> 531 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT
532 AATATTGAAA 5554
E--> 534 AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCTTTT
535 TTGCGGCATT 5614
E--> 537 TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG
538 CTGAAGATCA 5674
E--> 540 GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA
541 TCCTTGAGAG 5734
E--> 543 TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC
544 TATGTGGCGC 5794
E--> 546 GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC
547 ACTATTCTCA 5854
E--> 549 GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG
550 GCATGACAGT 5914
E--> 552 AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA
553 ACTTACTTCT 5974
E--> 555 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG
556 GGGATCATGT 6034
E--> 558 AACTCGCCTT GATCGTTGGG AACCAGGAGT GAATGAAGCC ATACCAAACG
559 ACGAGCGTGA 6094
E--> 561 CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCACAAA CTATTAAGTG
562 GCGAACTACT 6154
E--> 564 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG
565 TTGCAGGACC 6214
E--> 567 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG
568 GAGCCGGTGA 6274
E--> 570 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT
571 CCCGTATCGT 6334
E--> 573 AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC
574 AGATCGCTGA 6394
E--> 576 GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT
577 CATATATACT 6454
E--> 579 TTAGATTGAT TTAAACTTTC ATTTTAAATT TAAAAGGATC TAGGTGAAGA
580 TCCTTTTTGA 6514
E--> 582 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT
583 CAGACCCCGT 6574
E--> 585 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTTCTG CGCGTAATCT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

W--> 443      980                      985                      990
E--> 446 GTC AAA AGG GAA GGG AAC TCC AGT CAG AAC TGG CAA CGG TTT TAC CAA
W--> 447 4101
      448 Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Trp Gln Arg Phe Tyr Gln
W--> 449 995                      1000                      1005                      1010
E--> 451 CTG ACA AAG CTT CTG GAC TCC ATG CAT GAG GTG GTT GAG AAT CTC CTT
W--> 452 4149
      453 Leu Thr Lys Leu Leu Asp Ser Met His Glu Val Val Glu Asn Leu Leu
W--> 454                      1015                      1020                      1025
E--> 456 ACC TAC TGC TTC CAG ACA TTT TTG GAT AAG ACC ATG AGT ATT GAA TTC
W--> 457 4197
      458 Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Thr Met Ser Ile Glu Phe
W--> 459                      1030                      1035                      1040
E--> 461 CCA GAG ATG TTA GCT GAA ATC ATC ACT AAT CAG ATA CCA AAA TAT TCA
W--> 462 4245
      463 Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gln Ile Pro Lys Tyr Ser
W--> 464                      1045                      1050                      1055
E--> 466 AAT GGA AAT ATC AAA AAG CTT CTG TTT CAT CAA AAA TGA CTGCCTTACT
W--> 467 4294
      468 Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gln Lys *
W--> 469                      1060                      1065                      1070
E--> 471 AAGAAAGGTT GCCTTAAAGA AAGTTGAATT TATAGTCTAG AGTCGACCCG
      472 GGCGGCCGCT 4354
E--> 474 TCGAGCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAACATA
      475 GAATGCAGTG 4414
E--> 477 AAAAAAATGC TTTATTTGTG AAATTTGTGA TGCTATTGCT TTATTTGTAA
      478 CCATTATAAG 4474
E--> 480 CTGCAATAAA CAAGTTAACA ACAACAATTG CATTCATTTT ATGTTTCAGG
      481 TTCAGGGGGA 4534
E--> 483 GATGTGGGAG GTTTTAA GCAAGTAAAA CCTCTACAAA TGTGGTAAAA
      484 TCGATAAGGA 4594
E--> 486 TCCGGGCTGG CGTAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC
      487 AGTTGCGCAG 4654
E--> 489 CCTGAATGGC GAATGGACGC GCCCTGTAGC GCGGCATTAA GCGCGGCGGG
      490 TGTGGTGGTT 4714
E--> 492 ACGCGCAGCG TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT
      493 CGCTTTCTTC 4774
E--> 495 CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCCTCAAG CTCTAAATCG
      496 GGGGCTCCCT 4834
E--> 498 TTAGGGTTCC GATTTAGAGC TTTACGGCAC CTCGACCGCA AAAAATTGA
      499 TTTGGGTGAT 4894
E--> 501 GGTTACGTA GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC
      502 GTTGGAGTCC 4954
E--> 504 ACGTCTTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC
      505 TATCTCGGTC 5014
E--> 507 TATCTTTTG ATTTATAAGG GATTTTGCCG ATTTGCGCCT ATTGGTTAAA
      508 AAATGAGCTG 5074
E--> 510 ATTTAACAAA TATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT
      511 TTCGCCTGAT 5134

```

same

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002
TIME: 14:48:40

Input Set : A:\W065457.txt
Output Set: N:\CRF3\03252002\J001486.raw

586 GCTGCTTGCA 6634
E--> 588 AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC
589 TACCAACTCT 6694
E--> 591 TTTTCCGAAG GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC
592 TTCTAGTGTA 6754
E--> 594 GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC
595 TCGCTCTGCT 6814
E--> 597 AATCCTGTGA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG
598 GGTTGGACTC 6874
E--> 600 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT
601 CGTGCACACA 6934
E--> 603 GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG
604 AGCTATGAGA 6994
E--> 606 AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG
607 GCAGGGTCGG 7054
E--> 609 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT
610 ATAGTCCTGT 7114
E--> 612 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG
613 GGGGGCGGAG 7174
E--> 615 CCTATGGAAA AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCTTTT
616 GCTGGCCTTT 7234
618 TGCTCACATG GCTCGACAGA TCT

621 (2) INFORMATION FOR SEQ ID NO: 2:

623 (i) SEQUENCE CHARACTERISTICS:

624 (A) LENGTH: 1071 amino acids

625 (B) TYPE: amino acid

626 (D) TOPOLOGY: linear

628 (ii) MOLECULE TYPE: protein

630 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

632 Met Ala His His His His His Gly Tyr Pro Tyr Asp Val Pro Asp
633 1 5 10 15
635 Tyr Ala Gln Ser Ser Ala Met Ser Lys Gly Glu Glu Leu Phe Thr Gly
636 20 25 30
638 Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
639 35 40 45
641 Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
642 50 55 60
644 Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
645 65 70 75 80
647 Thr Leu Val Thr Thr Phe Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr
648 85 90 95
650 Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu
651 100 105 110
653 Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr
654 115 120 125
656 Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg
657 130 135 140
659 Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly
660 145 150 155 160

7257

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

662 His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala
663      165      170      175
665 Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn
666      180      185      190
668 Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr
669      195      200      205
671 Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser
672      210      215      220
674 Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met
675 225      230      235      240
677 Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp
678      245      250      255
680 Glu Leu Tyr Lys Gly Ala Gly Ala Gly Ala Gly Ala Gly Ala Ile Ser
681      260      265      270
683 Ala Leu Ile Leu Asp Ser Lys Glu Ser Leu Ala Pro Pro Gly Arg Asp
684      275      280      285
686 Glu Val Pro Gly Ser Leu Leu Gly Gln Gly Arg Gly Ser Val Met Asp
687      290      295      300
689 Phe Tyr Lys Ser Leu Arg Gly Gly Ala Thr Val Lys Val Ser Ala Ser
690 305      310      315      320
692 Ser Pro Ser Val Ala Ala Ala Ser Gln Ala Asp Ser Lys Gln Gln Arg
693      325      330      335
695 Ile Leu Leu Asp Phe Ser Lys Gly Ser Thr Ser Asn Val Gln Gln Arg
696      340      345      350
698 Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
699      355      360      365
701 Gln Gln Gln Pro Gly Leu Ser Lys Ala Val Ser Leu Ser Met Gly Leu
702      370      375      380
704 Tyr Met Gly Glu Thr Glu Thr Lys Val Met Gly Asn Asp Leu Gly Tyr
705 385      390      395      400
707 Pro Gln Gln Gly Gln Leu Gly Leu Ser Ser Gly Glu Thr Asp Phe Arg
708      405      410      415
710 Leu Leu Glu Glu Ser Ile Ala Asn Leu Asn Arg Ser Thr Ser Val Pro
711      420      425      430
713 Glu Asn Pro Lys Ser Ser Thr Ser Ala Thr Gly Cys Ala Thr Pro Thr
714      435      440      445
717 Glu Lys Glu Phe Pro Lys Thr His Ser Asp Ala Ser Ser Glu Gln Gln
718      450      455      460
720 Asn Arg Lys Ser Gln Thr Gly Thr Asn Gly Gly Ser Val Lys Leu Tyr
721 465      470      475      480
723 Pro Thr Asp Gln Ser Thr Phe Asp Leu Leu Lys Asp Leu Glu Phe Ser
E--> 724      485->      490->      495->
726 Ala Gly Ser Pro Ser Lys Asp Thr Asn Glu Ser Pro Trp Arg Ser Asp
E--> 727      500      505      510
729 Leu Leu Ile Asp Glu Asn Leu Leu Ser Pro Leu Ala Gly Glu Asp Asp
E 730      515      520      525
732 Pro Phe Leu Leu Glu Gly Asn Thr Asn Glu Asp Cys Lys Pro Leu Ile
E--> 733      530      535      540
735 Leu Pro Asp Thr Lys Pro Lys Ile Lys Asp Thr Gly Asp Thr Ile Leu

```

(see item 3
on Error Summary
sheet)

move these numbers
directly
under respective
amino acids

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

```

E--> 736 545          550          555          560
      738 Ser Ser Pro Ser Ser Val Ala Leu Pro Gln Val Lys Thr Glu Lys Asp
E--> 739          565          570          575
      741 Asp Phe Ile Glu Leu Cys Thr Pro Gly Val Ile Lys Gln Glu Lys Leu
E--> 742          580          585          590
      744 Gly Pro Val Tyr Cys Gln Ala Ser Phe Ser Gly Thr Asn Ile Ile Gly
E--> 745          595          600          605
      747 Asn Lys Met Ser Ala Ile Ser Val His Gly Val Ser Thr Ser Gly Gly
E--> 748          610          615          620
      750 Gln Met Tyr His Tyr Asp Met Asn Thr Ala Ser Leu Ser Gln Gln Gln
E--> 751 625          630          635          640
      753 Asp Gln Lys Pro Val Phe Asn Val Ile Pro Pro Ile Pro Val Gly Ser
E--> 754          645          650          655
      756 Glu Asn Trp Asn Arg Cys Gln Gly Ser Gly Glu Asp Ser Leu Thr Ser
E--> 757          660          665          670
      759 Leu Gly Ala Leu Asn Phe Pro Gly Arg Ser Val Phe Ser Asn Gly Tyr
E--> 760          675          680          685
      762 Ser Ser Pro Gly Met Arg Pro Asp Val Ser Ser Pro Pro Ser Ser Ser
E--> 763          690          695          700
      765 Ser Ala Ala Thr Gly Pro Pro Pro Lys Leu Cys Leu Val Cys Ser Asp
E--> 766 705          710          715          720
      768 Glu Ala Ser Gly Cys His Tyr Gly Val Leu Thr Cys Gly Ser Cys Lys
E--> 769          725          730          735
      771 Val Phe Phe Lys Arg Ala Val Glu Gly Gln His Asn Tyr Leu Cys Ala
E--> 772          740          745          750
      775 Gly Arg Asn Asp Cys Ile Ile Asp Lys Ile Arg Arg Lys Asn Cys Pro
E--> 776          755          760          765
      778 Ala Cys Arg Tyr Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala
E--> 779          770          775          780
      781 Arg Lys Thr Lys Lys Lys Ile Lys Gly Ile Gln Gln Ala Thr Ala Gly
E--> 782 785          790          795          800
      784 Val Ser Gln Asp Thr Ser Glu Asn Pro Asn Lys Thr Ile Val Pro Ala
E--> 785          805          810          815
      787 Ala Leu Pro Gln Leu Thr Pro Thr Leu Val Ser Leu Leu Glu Val Ile
E--> 788          820          825          830
      790 Glu Pro Glu Val Leu Tyr Ala Gly Tyr Asp Ser Ser Val Pro Asp Ser
E--> 791          835          840          845
      793 Ala Trp Arg Ile Met Thr Thr Leu Asn Met Leu Gly Gly Arg Gln Val
E--> 794          850          855          860
      796 Ile Ala Ala Val Lys Trp Ala Lys Ala Ile Leu Gly Leu Arg Asn Leu
E--> 797 865          870          875          880
      799 His Leu Asp Asp Gln Met Thr Leu Leu Gln Tyr Ser Trp Met Phe Leu
E--> 800          885          890          895
      802 Met Ala Phe Ala Leu Gly Trp Arg Ser Tyr Arg Gln Ser Ser Gly Asn
E--> 803          900          905          910
      805 Leu Leu Cys Phe Ala Pro Asp Leu Ile Ile Asn Glu Gln Arg Met Ser
E--> 806          915          920          925
      808 Leu Pro Gly Met Tyr Asp Gln Cys Lys His Met Leu Phe Val Ser Ser
E--> 809          930          935          940

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002
TIME: 14:48:40

Input Set : A:\W065457.txt
Output Set: N:\CRF3\03252002\J001486.raw

811 Glu Leu Gln Arg Leu Val Ser Tyr Glu Glu Tyr Leu Cys Met Lys
E--> 812 945 950 955 960
814 Thr Leu Leu Leu Leu Ser Ser Val Pro Lys Glu Gly Leu Lys Ser Gln
E--> 815 965 970 975
817 Glu Leu Phe Asp Glu Ile Arg Met Thr Tyr Ile Lys Glu Leu Gly Lys
E--> 818 980 985 990
820 Ala Ile Val Lys Arg Glu Gly Asn Ser Ser Gln Asn Trp Gln Arg Phe
E--> 821 995 1000 1005
823 Tyr Gln Leu Thr Lys Leu Leu Asp Ser Met His Glu Val Val Glu Asn
E--> 824 1010 1015 1020
826 Leu Leu Thr Tyr Cys Phe Gln Thr Phe Leu Asp Lys Thr Met Ser Ile
E--> 827 1025 1030 1035 1040
829 Glu Phe Pro Glu Met Leu Ala Glu Ile Ile Thr Asn Gln Ile Pro Lys
E--> 830 1045 1050 1055
E--> 833 Tyr Ser Asn Gly Asn Ile Lys Lys Leu Leu Phe His Gln Lys *
E--> 834 1060 1065 1070

delete stop codon

885 (2) INFORMATION FOR SEQ ID NO: 6:

887 (i) SEQUENCE CHARACTERISTICS:

888 (A) LENGTH: 60 base pairs

889 (B) TYPE: nucleic acid

890 (C) STRANDEDNESS: single

891 (D) TOPOLOGY: linear

W--> 893 (ii) MOLECULE TYPE: oligonucleotide

896 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

E--> 898 TCGAGCGCGC AAGAACACAG TGTTCTGACG ACACGAAGAA CAGGATGTTC

899 TCGTACAGTG 60

902 (2) INFORMATION FOR SEQ ID NO: 7:

904 (i) SEQUENCE CHARACTERISTICS:

905 (A) LENGTH: 60 base pairs

906 (B) TYPE: nucleic acid

907 (C) STRANDEDNESS: single

908 (D) TOPOLOGY: linear

W--> 910 (ii) MOLECULE TYPE: oligonucleotide

913 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

E--> 915 TCGACACTGT ACGAGAACAT CCTGTTCTTC GTGTCGTCAG AACACTGTGT

916 TCTTGCGCGC 60

919 (2) INFORMATION FOR SEQ ID NO: 8:

921 (i) SEQUENCE CHARACTERISTICS:

922 (A) LENGTH: 60 base pairs

923 (B) TYPE: nucleic acid

924 (C) STRANDEDNESS: single

925 (D) TOPOLOGY: linear

W--> 927 (ii) MOLECULE TYPE: oligonucleotide

930 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

E--> 932 TCGAGCGCGC AAGGTCACAG TGACCTGACG ACACGAAGGT CAGGATGACC

933 TCGTACAGTG 60

936 (2) INFORMATION FOR SEQ ID NO: 9:

938 (i) SEQUENCE CHARACTERISTICS:

939 (A) LENGTH: 60 base pairs

*see item 1 on
Err
summary sheet*

same error

same

8/14

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:40

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

940 (B) TYPE: nucleic acid

941 (C) STRANDEDNESS: single

942 (D) TOPOLOGY: linear

944 (ii) MOLECULE TYPE: DNA (genomic)

947 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

E--> 949 TCGACACTGT ACGAGGTCAT CCTGACCTTC GTGTCGTCAG GTCACTGTGA

950 CCTTGCGCGC 60

same

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:41

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

L:34 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:35 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:38 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:59 M:220 C: Keyword misspelled or invalid format, [(D) TOPOLOGY:]
L:61 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:70 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:1
M:254 Repeated in SeqNo=1
L:127 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:129 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:132 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:134 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:138 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:140 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:143 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:145 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:148 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:150 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:153 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:155 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:158 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:160 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:163 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:165 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:168 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:170 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:173 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:175 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:178 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:180 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:183 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:185 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:188 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:190 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:193 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:195 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:198 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:200 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:203 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:205 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:210 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:212 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:215 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:217 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:220 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:222 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:227 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:230 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/001,486

DATE: 03/25/2002

TIME: 14:48:41

Input Set : A:\W065457.txt

Output Set: N:\CRF3\03252002\J001486.raw

L:232 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:235 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:237 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:240 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:242 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:245 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:247 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:250 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:252 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:724 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2
M:332 Repeated in SeqNo=2
L:833 M:342 E: Invalid Stop Code On Error, STOP CODON:*\nL:845 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3\nL:862 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4\nL:877 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5\nL:893 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6\nL:898 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:6\nL:910 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7\nL:915 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:7\nL:927 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8\nL:932 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:8\nL:949 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:9